

Genetica

Pesco, un genoma made in Italy

Davide Ederle
Laura Rossini
Ignazio Verde

La scelta della tecnica Sanger ha permesso di ottenere una sequenza di elevata qualità e di grande rilevanza per la ricerca in frutticoltura.

L'albero della pesca (*Prunus persica*), originario della Cina, è coltivato da almeno 4000 anni. È stato introdotto in Europa attraverso la Persia seguendo la via della seta nel primo secolo avanti Cristo [1]. Il sequenziamento del suo patrimonio genetico è stato promosso dall'Italia nel 2005 grazie all'iniziativa di Francesco Salamini¹. La spinta decisiva per portare a termine il lavoro si ebbe però nel 2008, quando, attorno al progetto italiano *Drupomics*, coordinato dal Consiglio per la ricerca e la sperimentazione in agricoltura (Cra) e finanziato dal Ministero per le politiche agricole, alimentari e forestali (Mipaaf) con 2,58 milioni di euro, si costituì un consorzio internazionale *The international peach genome initiative* (Ipgi) a guida congiunta Italia-Usa. A sancire la paternità *made in Italy* del progetto, che vede la partecipazione anche di cileni, americani, spagnoli e francesi, è l'attività di coordinamento del consorzio con il Cra di Roma, il Parco tecnologico padano di Lodi, l'Istituto di genomica applicata e l'Università degli Studi di Udine, affiancati da ricercatori americani. A rendere interessante il genoma del pesco è il fatto che, con 1,6 milioni di tonnellate all'anno, l'Italia è il secondo produttore mondiale dopo la Cina. Inoltre, il pesco rappresenta una specie modello per le piante arboree da frutto e per l'intera famiglia delle *Rosaceae*, che comprende molte altre specie coltivate quali il melo, la fragola, il lampone, il ciliegio, l'albicocco. Il genoma del pesco,

composto da poco meno di 230 milioni di nucleotidi (Mb), è più piccolo rispetto a quello di altre *Rosaceae* (per esempio il melo circa 700 Mb). Ulteriori caratteristiche che rendono la specie ideale per lo studio sono il breve periodo giovanile, ridotto quasi a una sola vegetazione, la riproduzione autogama, per cui non richiede varietà impollinatrici (l'impollinazione è entomofila, operata dalle api e da altri insetti) e l'ampia disponibilità di materiali come mappe genetiche e fisiche e individui doppi aploidi, ossia completamente omozigoti.

Metodi

Il sequenziamento del pesco è stato condotto su un individuo doppio aploide della cultivar 'Lovell' (PLov2-2N), omozigote a tutti i loci. Come metodologia di sequenziamento è stata adottata la tecnica Sanger, costosa ma capace di ottenere sequenze di maggiore qualità rispetto alle metodiche di nuova generazione. Il sequenziamento Sanger fornisce, infatti, vantaggi in termini di accuratezza e di lunghezza delle sequenze (800 vs 100-500 bp a seconda del metodo). Le sequenze sono state assemblate con il software Arachne [3] e successivamente ricondotte alla loro collocazione cromosomica mediante l'utilizzo di diverse mappe genetiche [4, 5, 6, 7]. L'analisi della variabilità intra e interspecifica è stata condotta su diverse accessioni di pesco e specie affini utilizzando un metodo di sequenziamento massivo sulla piattaforma Illumina. L'iniziale annotazione funzionale del genoma è stata condotta con strumenti bioinformatici e successivamente raffinata manualmente per un sotto-set di geni candidati per i processi di maturazione e i caratteri di qualità del frutto.

¹ Francesco Salamini è stato professore di Botanica e Fisiologia presso l'Università Cattolica del Sacro cuore di Piacenza e Direttore della Sezione maiscoltura di Bergamo dell'Istituto sperimentale per la cerealicoltura di Roma. Dall'11 settembre 2009 è Presidente dell'Istituto agrario di San Michele all'Adige.

La sequenza

La sequenza ottenuta è risultata di alta qualità (99,96% di accuratezza), poco frammentata (i 230 milioni di basi del Dna del pesco, un terzo di quelle del melo e un terzo di quelle dell'uomo, sono racchiusi in 202 macromolecole o *scaffold*), completamente ancorata ai cromosomi (il 99,3% della sequenza nella versione aggiornata ha una collocazione sugli 8 cromosomi della specie) e orientata (97,9%). Scelte cruciali ai fini della qualità della sequenza sono state l'utilizzo di un individuo completamente omozigote (questo consente di annullare la diversità allelica intra individuo che rende difficoltose le operazioni di assemblaggio) e l'adozione della metodologia Sanger.

Gli elementi genetici

Grazie a questo studio sono stati identificati 27.852 geni, un numero molto simile a quello trovato per il genoma umano. Di questi, 672 geni – coinvolti nel controllo di importanti caratteri di qualità (maturazione, aroma, contenuto zuccherino e in pigmenti, forma della pianta e del frutto) – sono stati studiati in profondità e hanno consentito di mettere in evidenza alcune peculiarità delle drupacee (formazione del nocciolo) e delle rosacee (metabolismo del sorbitolo). Inoltre, sono state individuate le sequenze ripetute e mobili (trasposoni), che costituiscono il 29% circa del genoma della specie, e diversi Rna non codificanti come, per esempio, i microRna.

Diversità genetica e domesticazione

Sono state inoltre individuate sul Dna un milione di varianti genetiche (marcatori molecolari *Single nucleotide polymorphism*, Snp) che hanno consentito di ricostruire la storia evolutiva della specie. Il confronto della variabilità tra accessioni orientali e occidentali di pesco e specie selvatiche affini ha permesso di evidenziare infatti due momenti principali per la diversificazione del Dna attuale della pianta, eventi che hanno portato a una perdita considerevole della biodiversità (colli di bottiglia, *bottlenecks*): il primo legato alla domesticazione originaria, avvenuta in Cina circa 4-5000 anni fa, seguita da una seconda perdita di variabilità durante la dispersione verso occidente attraverso la Persia e la recente e intensa attività di miglioramento genetico (*breeding*). Questa perdita di variabilità ha portato alla fissazione di aplotipi (una serie di alleli trasmessi in blocco) che si estendono per lunghi tratti cromosomici.

L'evoluzione del genoma

L'analisi comparativa con altre specie sequenziate (vite, melo, pioppo) ha messo in evidenza che il pesco condivide con le altre dicotiledoni una esaploidizzazione antica del genoma avvenuta in un comune progenitore, causata dalla fusione di tre genomi ancestrali e seguita dalla loro diploidizzazione (duplicazione del numero cromosomico) avvenuta in tre diversi cicli occorsi centinaia di milioni di anni fa (così come evidenziato nella vite [8]). Al contrario di altre specie come il melo [9] e il pioppo [10], il pesco non ha subito duplicazioni recenti del genoma. Il genoma del pesco (e delle *Rosaceae*), contrariamente a quello della vite, è andato incontro invece a successivi riarrangiamenti del materiale genetico come fusioni e rotture tra i cromosomi derivanti dall'ancestrale comune delle dicotiledoni.

I geni e i marcatori molecolari individuati renderanno possibile la selezione di varietà migliorate nelle loro caratteristiche qualitative e nelle proprietà nutraceutiche, nonché dotate di maggiore adattabilità ai cambiamenti climatici e resistenza ai parassiti. Il tutto con minori costi e in tempi più brevi.

Un aspetto interessante delle specie appartenenti alla famiglia delle *Rosaceae*, di cui il pesco fa parte, è quello di associare a un genoma relativamente piccolo una grande variabilità di forme sia nell'habitus vegetativo (si passa dalla forma erbacea, alla cespugliosa, arbustiva e arborea), sia nel tipo di frutto (achenii della fragola, pomi del melo, drupe del pesco e bacche del lampone). Ora che si dispone del genoma non solo del pesco [2], ma anche del melo [9], del pero [11] e della fragola [12] è possibile capire quali siano i geni che controllano questi caratteri. Sarà inoltre possibile rinviare la ridotta variabilità della specie coltivata attraverso l'introgresione mirata di caratteri genetici importanti come, per esempio, la resistenza alle malattie, presenti nelle specie selvatiche affini (per esempio *P. davidiana* e *P. kansuensis*) e di cui la specie coltivata è carente.

La sequenza, grazie alla sua accuratezza e completezza, potrà fungere da valido riferimento per la ricostruzione dei genomi di altre importanti specie sia all'interno che all'esterno della famiglia. Per esempio, studi preliminari hanno consentito di utilizzare la sequenza del pesco come riferimento per la ricostruzione di genomi di specie al di fuori della famiglia (*Cannabis*,

<http://genomevolution.org/wiki/index.php/Pseudoassembly>).

Il genoma del pesco è accessibile e scaricabile ai seguenti indirizzi internet:

www.phytozome.net/peach

http://www.rosaceae.org/species/prunus_persica/genome_v1.0

http://services.appliedgenomics.org/fgb2/iga/prunus_public/gbrowse/prunus_public/

Riferimenti bibliografici

[1] Faust M., Timon B., 1995. Origin and dissemination of peach. *Horticulture oriented to recreation and technique*. Reviews 17, 331-379.

[2] Verde I. et al., 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature genetics*, 45, 487-494. Doi: 10.1038/ng. 2586
<http://www.nature.com/ng/journal/v45/n5/full/ng.2586.html>

[3] Jaffe D. B. et al., 2003. Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome research*, 13, 91-96.

[4] Dirlewanger E. et al., 2004. Comparative mapping and marker-assisted selection in *Rosaceae* fruit crops. *Proceedings of the national academy of sciences*, 101, 9891-9896. Usa.

[5] Howad W. et al., 2005. Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics*, 171, 1305-1309.

[6] Verde I. et al., 2005. Microsatellite and Aflp markers in the *Prunus persica* [L.(Batsch)] x *P. ferganensis* BC1 linkage map: saturation and coverage improvement. *Theoretical and applied genetics*, 111, 1013-1021.

[7] Eduardo I. et al., 2011. QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. *Tree genetics et genomes*, 7, 323-335.

[8] Jaillon O. et al., 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449, 463-467.

[9] Velasco R. et al., 2010. The genome of the domesticated apple (*Malus x domestica* Borkh.). *Nature genetics*, 42, 833-839.

[10] Tuskan G. A. et al., 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313, 1596-1604.

[11] Wu J. et al., 2013. The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome research*, 23 (2), 396-408.

[12] Shulaev V. et al, 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nature genetics*, 43, 109-116.



Davide Ederle è biotecnologo presso il Parco tecnologico padano di Lodi.

Laura Rossini è ricercatrice confermata presso il Dipartimento di Produzione vegetale dell'Università degli Studi di Milano.

Ignazio Verde è ricercatore presso il Consiglio per la ricerca e la sperimentazione in agricoltura - Centro di ricerca per la frutticoltura di Roma.

www.intersezioni.eu



Regione Lombardia

Fondo europeo agricolo per lo sviluppo rurale: l'Europa investe nelle zone rurali
PSR 2007-2013 – Direzione Generale Agricoltura